Time Series and Forecasting Assignment

Code

1.Time Series Model

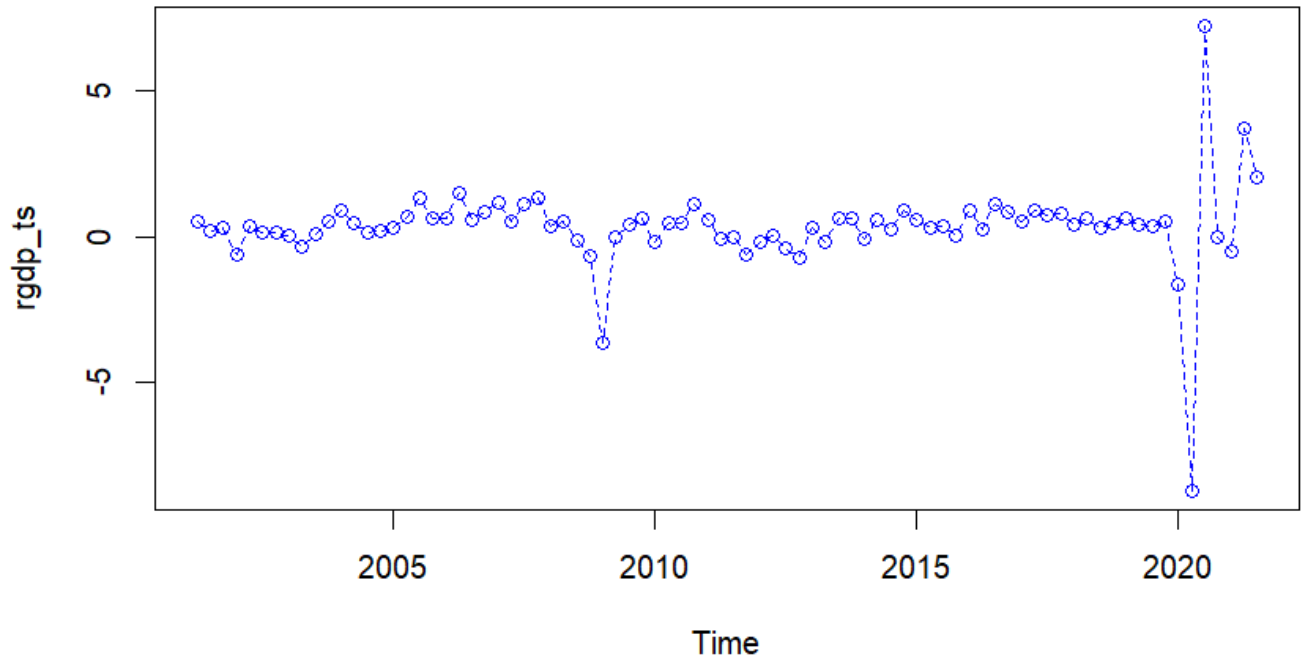
Hide

```
rm(list = ls())
library(tseries)
library(fGarch)
library(rugarch)
library(Hmisc)
library(readxl)
excel_file <- "EULN1NETH.xlsx"
df <- read_excel(excel_file)
```

```
New names:
```

Hide

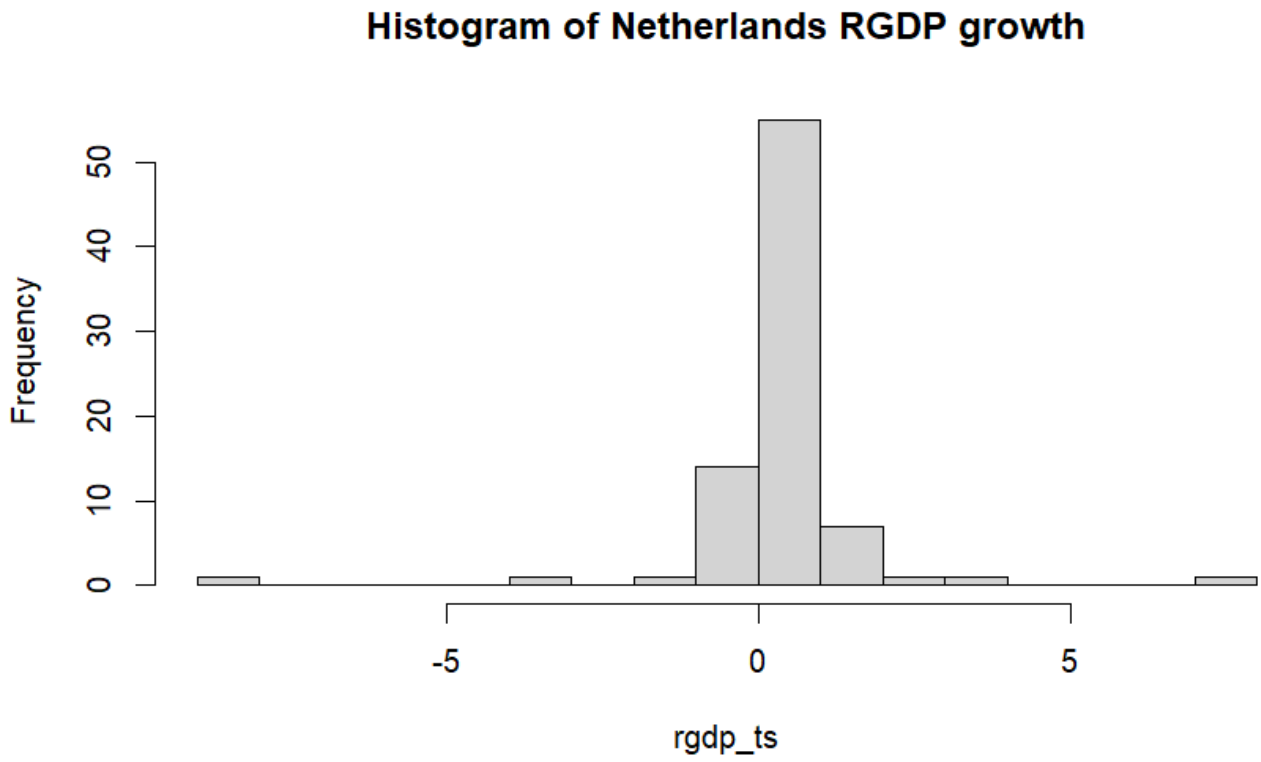
```
df <- df[-c(1:10), ]
df <- df[, -1]
df$x1 <- as.double(df$x1)
df$x2 <- as.double(df$x2)
df$x3 <- as.double(df$x3)
df$x4 <- as.double(df$x4)
df$x5 <- as.double(df$x5)
df$x6 <- as.double(df$x6)
df$x7 <- as.double(df$x7)
df$x8 <- as.double(df$x8)
df$x9 <- as.double(df$x9)
df$x10 <- as.double(df$x10)
df$x11<- as.double(df$x11)
df$x12 <- as.double(df$x12)
df$x13 <- as.double(df$x13)
df$x14 <- as.double(df$x14)
df$x15 <- as.double(df$x15)
df$x16 <- as.double(df$x16)
df$y <- as.double(df$y)
rgdp_ts =ts(df$y, frequency=4, start = c(2001,2))
plot(rgdp_ts, type="o", col="blue", lty="dashed")
```



We will plot a histogram of the distribution of the RGDP growth values.

Hide

```
hist(rgdp_ts, nclass=20, main="Histogram of Netherlands RGDP growth")
```

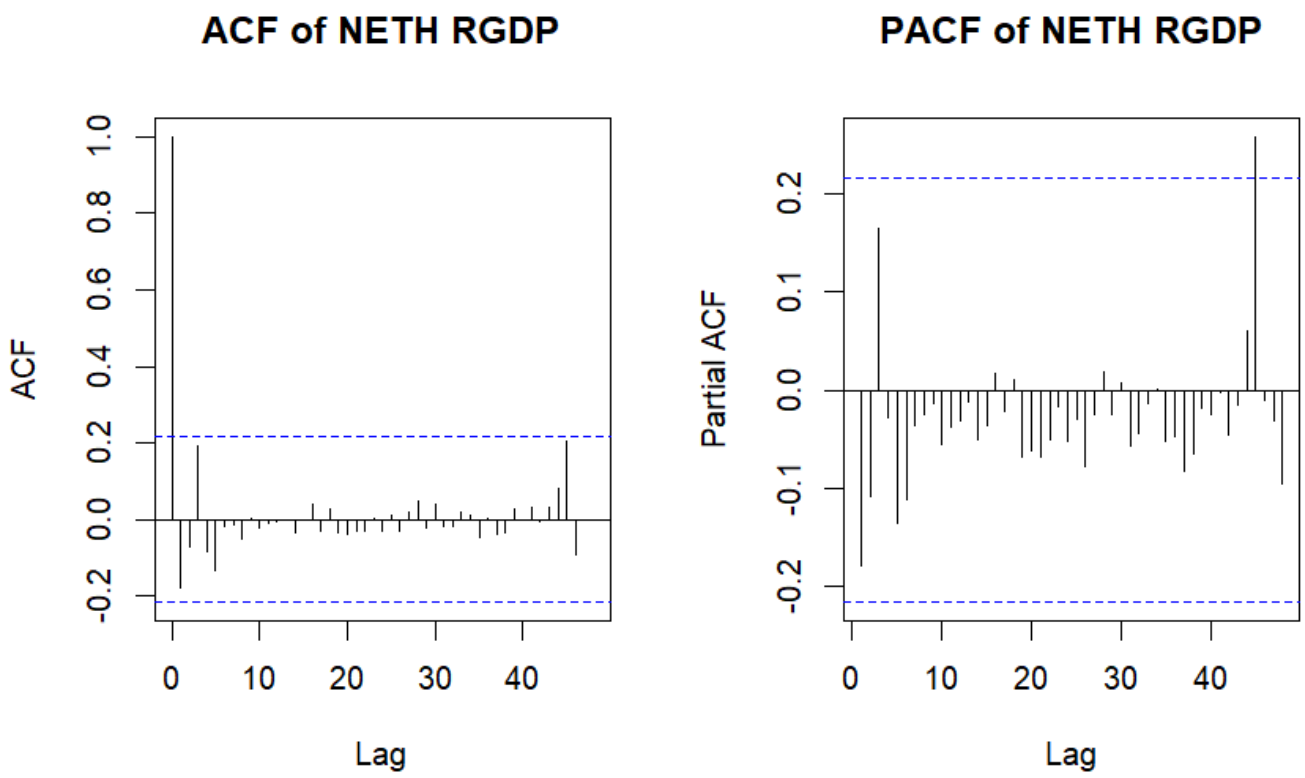


Bellow we will plot the ACF and PACF plots to find the autocorrelation and partial autocorrelation per lag(1-48) for our data. We can see that for all lags, the autocorrelations and partial autocorrelations are below the critical value line, which means that they are not significant and there is no autocorrelation between the yt values

(PACF) or the error values (ACF). On lag 45 we can see that the partial autocorrelation goes over the critical value line. We will consider that as noise, as 45 periods before can't be significant to the current values.

Hide

```
par(mfrow=c(1,2))        # set up the graphics
acf(ts(rgdp_ts,freq=1), 48, main="ACF of NETH RGDP")        # autocorrelation function plot
pacf(ts(rgdp_ts,freq=1), 48, main="PACF of NETH RGDP")    # partial autocorrelation function
```



Hide

```
m=ar(rgdp_ts)
m
```

```

Call:
ar(x = rgdp_ts)

Coefficients:
      1
-0.1785

Order selected 1  sigma^2 estimated as  2.191
```

Hide

```
m$order
```

```
[1] 1
```

Below we run the Augmented Dickey-Fuller Test to see if our data are stationary, getting a p-value smaller than 0.01. We reject the null hypothesis that the data are not stationary.

Hide

```
adf.test(rgdp_ts,k=1)
```

```
Warning: p-value smaller than printed p-value
```

```

    Augmented Dickey-Fuller Test

data:  rgdp_ts
Dickey-Fuller = -7.2574, Lag order = 1, p-value = 0.01
alternative hypothesis: stationary
```

Bellow we run both Box-Pierce and Ljung-box tests to see if what we saw on the plots before is supported. The results agree that there is no significant autocorrelation between our data for any lag. (both p-values close to 1, we don't reject the null hypothesis, that there is no autocorrelation between our data). Because of these results we cannot continue with the time series modeling of our data, as there is no connection (correlation) between the values and we cannot extract any information to apply to our model. The values are simply independent and we cannot model any relationship between them. We will continue with the regression analysis.

Hide

```
res1=Box.test(rgdp_ts,48,type="Box-Pierce")
res2=Box.test(rgdp_ts,48,type="Ljung-Box")
res1
```

```

    Box-Pierce test

data:  rgdp_ts
X-squared = 14.972, df = 48, p-value = 1
```

Hide

```
res2
```

```

    Box-Ljung test

data:  rgdp_ts
X-squared = 22.522, df = 48, p-value = 0.9994
```

Hide

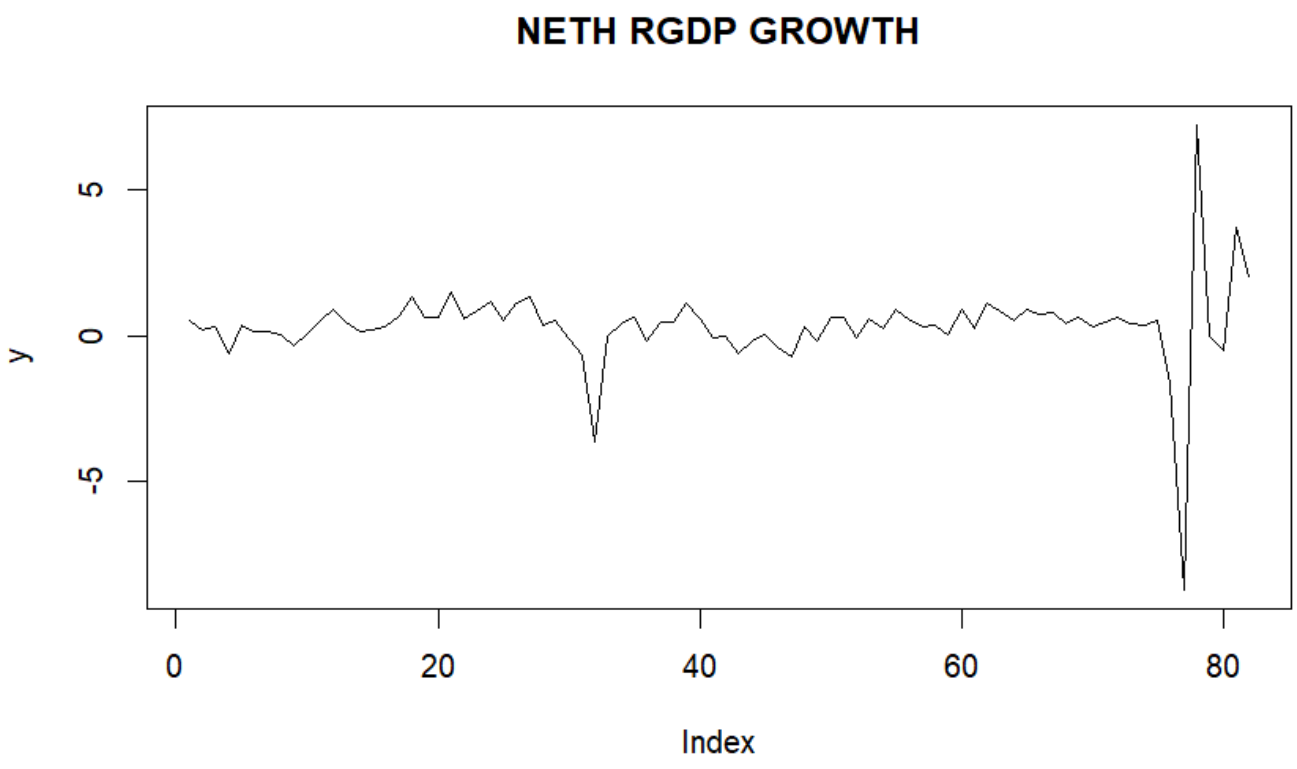
```
# END OF TIME SERIES MODELING: NO AUTOCORRELATION IN RESIDUALS
```

2. Combination of Approaches

We will plot our data again.

Hide

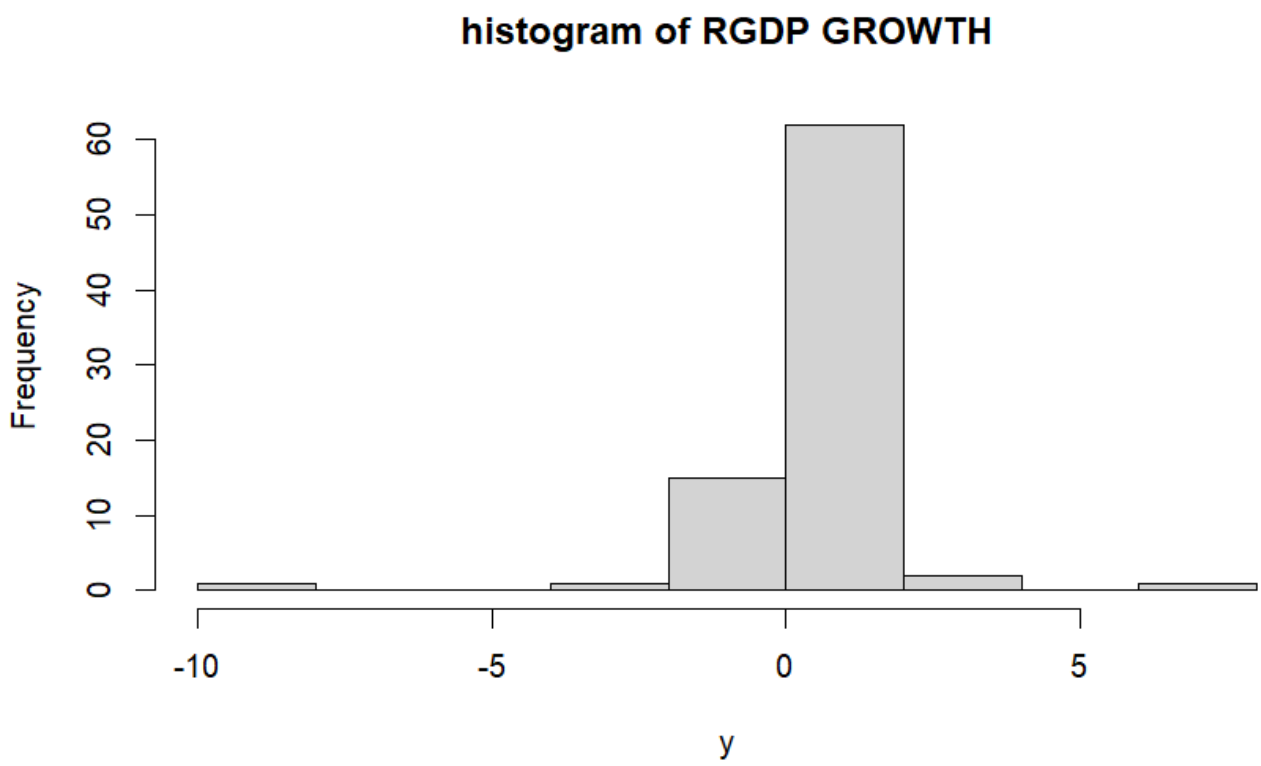
```
y <- df$y
x1 <- df$x1
x2 <- df$x2
x3 <- df$x3
x4 <- df$x4
x5 <- df$x5
x6 <- df$x6
x7 <- df$x7
x8 <- df$x8
x9 <- df$x9
x10 <- df$x10
x11 <- df$x11
x12 <- df$x12
x13 <- df$x13
x14 <- df$x14
x15 <- df$x15
x16 <- df$x16


# Summary Statistics and plots
plot(y, type="l", main="NETH RGDP GROWTH")
```



Bellow we plot the distribution of RGDP Growth. We can see that our data do not follow the normal distribution. (Tests will prove below).

Hide

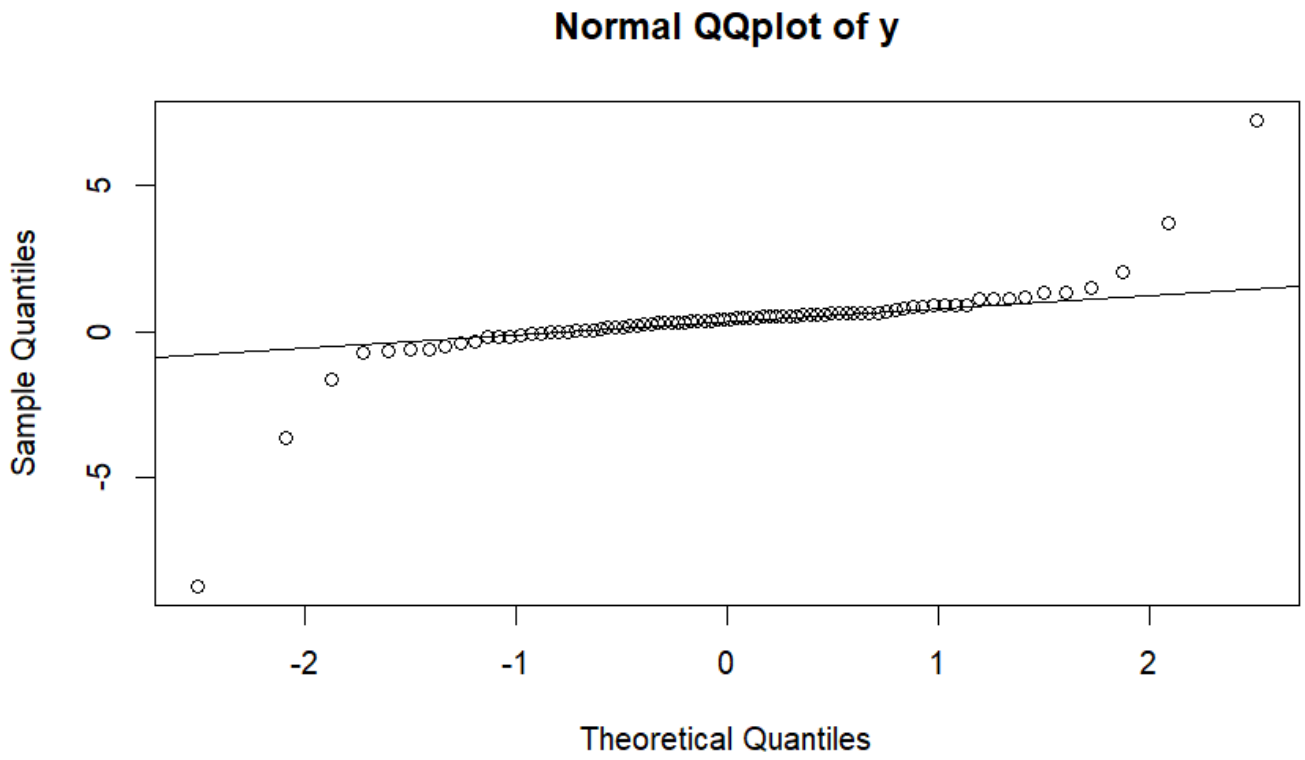
```
hist(y, main="histogram of RGDP GROWTH")
```



The QQ plot for the normal distribution below shows us that our data are not normal for the tail values. The tails are fatter and the data probably follow a t distribution.

Hide

```
qqnorm(y,main="Normal QQplot of y") # normal Q-Q plot
qqline(y)
```



We will also run normality tests that will show us that the null hypothesis (normality) is rejected.

Hide

```
jarque.bera.test(y)
```

```
    Jarque Bera Test

data:  y
X-squared = 1505.4, df = 2, p-value < 2.2e-16
```

Hide

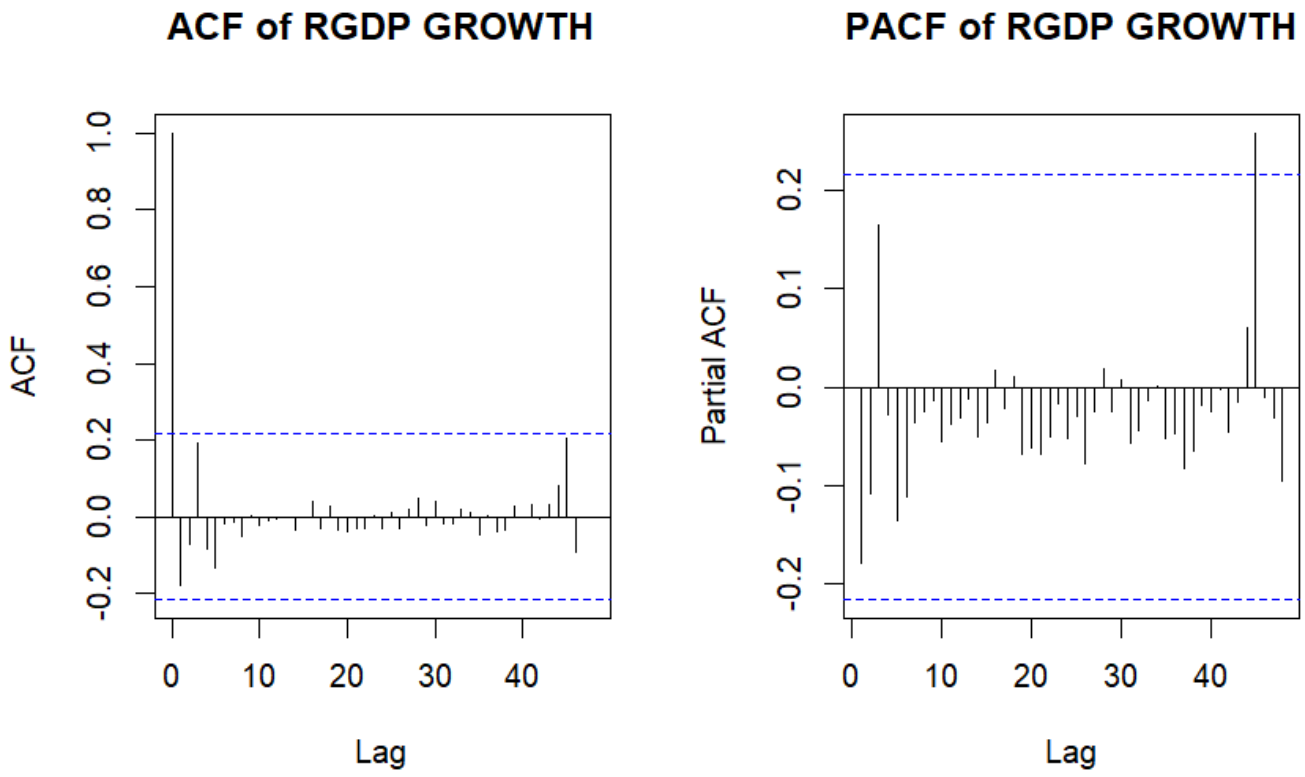
```
shapiro.test(y)
```

```
    Shapiro-Wilk normality test

data:  y
W = 0.59505, p-value = 1.07e-13
```

We can run agai both ACF and PACF plots to get the same results we discussed before.

Hide

```
par(mfrow=c(1,2))
acf(y, 48, main="ACF of RGDP GROWTH")
pacf(y, 48, main="PACF of RGDP GROWTH")
```



Hide

```
Box.test(y,lag=12,type="Ljung")
```

```
    Box-Ljung test

data:  y
X-squared = 8.9302, df = 12, p-value = 0.7089
```

To create a regression model we want to see the correlation between each xi attributes of our data. Bellow we can see a plot that shows the pairwise relationship of the attributes of our data.

Hide

```
# Correlation coefficients
cor(cbind(y,x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12,x13,x14,x15,x16))
```

```
                y           x1           x2          x3           x4          x5           x6
x7           x8           x9         x10
y    1.0000000000 -0.180077152 -0.074862196  0.19947778 -0.086364677  0.47880210  0.15400413
0.0007699958  0.602210748  0.261731794 -0.20051283
x1  -0.1800771521  1.000000000 -0.221452036 -0.06828086  0.202702451 -0.11585394  0.01680443
-0.5793620335  0.005818457 -0.112118128 -0.23054216
x2  -0.0748621957 -0.221452036  1.000000000 -0.20902639 -0.059339326  0.02286495 -0.02000007
0.2106256388 -0.351825161 -0.007949768  0.03509638
x3   0.1994777839 -0.068280859 -0.209026389  1.00000000 -0.207831672 -0.13248539  0.20223990
0.1549182365  0.239737751 -0.011321594 -0.31368963
x4  -0.0863646770  0.202702451 -0.059339326 -0.20783167  1.000000000 -0.12421314 -0.03927852
-0.0949824396 -0.225792399  0.065229599 -0.10440078
x5   0.4788020994 -0.115853943  0.022864955 -0.13248539 -0.124213143  1.00000000 -0.08545956
-0.1216743879  0.367421381  0.057372398 -0.07976293
x6   0.1540041293  0.016804427 -0.020000070  0.20223990 -0.039278523 -0.08545956  1.00000000
0.0565822430  0.121764644 -0.084243242 -0.06859260
x7   0.0007699958 -0.579362033  0.210625639  0.15491824 -0.094982440 -0.12167439  0.05658224
1.0000000000 -0.172725713  0.073792209  0.08330006
x8   0.6022107485  0.005818457 -0.351825161  0.23973775 -0.225792399  0.36742138  0.12176464
-0.1727257130  1.000000000  0.167482680 -0.15220426
x9   0.2617317937 -0.112118128 -0.007949768 -0.01132159  0.065229599  0.05737240 -0.08424324
0.0737922088  0.167482680  1.000000000 -0.45456598
x10 -0.2005128311 -0.230542161  0.035096380 -0.31368963 -0.104400779 -0.07976293 -0.06859260
0.0833000585 -0.152204263 -0.454565976  1.00000000
x11  0.1433240139  0.088607983  0.271179960 -0.03964014  0.002916711  0.24401379 -0.01364589
-0.2400047369 -0.010839063  0.005971451 -0.10203878
x12  0.2433065555 -0.226973807  0.063292357 -0.21795749 -0.119449802  0.66002011  0.07526716
-0.0018910901  0.135870601 -0.102397358  0.05585631
x13 -0.0999217311  0.133884781  0.148376607 -0.11240766 -0.084982839 -0.14667085 -0.09784846
0.0602280382 -0.052793727  0.116645669  0.13316748
x14  0.0687797334  0.028180613  0.161589514  0.05461283 -0.109943601  0.35614501 -0.22316188
-0.0255780360  0.157260943  0.015962616 -0.16975311
x15  0.2636131633  0.014309363 -0.022587635 -0.17318565 -0.150710470  0.48332151 -0.02723455
-0.1029782547  0.317699461  0.009488312  0.04032444
x16  0.4800519396  0.285792441  0.141067558  0.14689513 -0.300419929  0.32610453  0.03728885
-0.1723616072  0.354074501  0.077014152 -0.19934871
             x11         x12           x13         x14          x15           x16
y    0.143324014  0.24330656 -0.0999217311  0.06877973  0.263613163  0.4800519396
x1   0.088607983 -0.22697381  0.1338847806  0.02818061  0.014309363  0.2857924415
x2   0.271179960  0.06329236  0.1483766067  0.16158951 -0.022587635  0.1410675585
x3  -0.039640142 -0.21795749 -0.1124076554  0.05461283 -0.173185654  0.1468951266
x4   0.002916711 -0.11944980 -0.0849828387 -0.10994360 -0.150710470 -0.3004199294
x5   0.244013789  0.66002011 -0.1466708526  0.35614501  0.483321507  0.3261045284
x6  -0.013645886  0.07526716 -0.0978484626 -0.22316188 -0.027234546  0.0372888502
x7  -0.240004737 -0.00189109  0.0602280382 -0.02557804 -0.102978255 -0.1723616072
x8  -0.010839063  0.13587060 -0.0527937268  0.15726094  0.317699461  0.3540745007
x9   0.005971451 -0.10239736  0.1166456695  0.01596262  0.009488312  0.0770141524
x10 -0.102038781  0.05585631  0.1331674828 -0.16975311  0.040324439 -0.1993487132
x11  1.000000000  0.13443531 -0.0051974485  0.20485436  0.068275568  0.2158021291
x12  0.134435313  1.00000000  0.1466874982  0.35378889  0.494853521  0.0705706311
x13 -0.005197448  0.14668750  1.0000000000  0.16715678  0.057448870 -0.0004810736
x14  0.204854364  0.35378889  0.1671567756  1.00000000  0.367730883  0.1735265207
x15  0.068275568  0.49485352  0.0574488702  0.36773088  1.000000000  0.3893019539
x16  0.215802129  0.07057063 -0.0004810736  0.17352652  0.389301954  1.0000000000
```

Hide

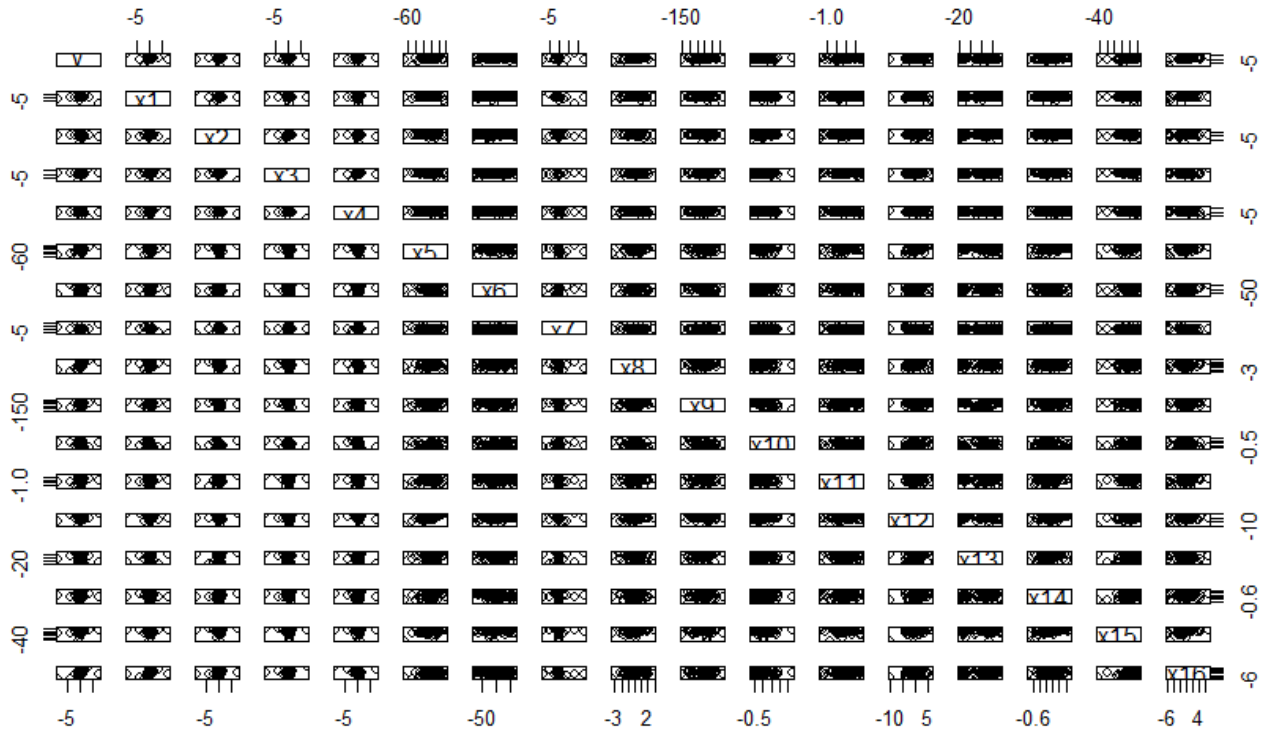
```
# Correlation coefficients and p-values
rcorr(as.matrix(cbind(y,x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12,x13,x14,x15,x16)))
```

```
        y    x1    x2    x3    x4    x5    x6    x7    x8    x9   x10   x11   x12   x13   x14
x15   x16
y    1.00 -0.18 -0.07  0.20 -0.09  0.48  0.15  0.00  0.60  0.26 -0.20  0.14  0.24 -0.10  0.07
0.26  0.48
x1  -0.18  1.00 -0.22 -0.07  0.20 -0.12  0.02 -0.58  0.01 -0.11 -0.23  0.09 -0.23  0.13  0.03
0.01  0.29
x2  -0.07 -0.22  1.00 -0.21 -0.06  0.02 -0.02  0.21 -0.35 -0.01  0.04  0.27  0.06  0.15  0.16
-0.02  0.14
x3   0.20 -0.07 -0.21  1.00 -0.21 -0.13  0.20  0.15  0.24 -0.01 -0.31 -0.04 -0.22 -0.11  0.05
-0.17  0.15
x4  -0.09  0.20 -0.06 -0.21  1.00 -0.12 -0.04 -0.09 -0.23  0.07 -0.10  0.00 -0.12 -0.08 -0.11
-0.15 -0.30
x5   0.48 -0.12  0.02 -0.13 -0.12  1.00 -0.09 -0.12  0.37  0.06 -0.08  0.24  0.66 -0.15  0.36
0.48  0.33
x6   0.15  0.02 -0.02  0.20 -0.04 -0.09  1.00  0.06  0.12 -0.08 -0.07 -0.01  0.08 -0.10 -0.22
-0.03  0.04
x7   0.00 -0.58  0.21  0.15 -0.09 -0.12  0.06  1.00 -0.17  0.07  0.08 -0.24  0.00  0.06 -0.03
-0.10 -0.17
x8   0.60  0.01 -0.35  0.24 -0.23  0.37  0.12 -0.17  1.00  0.17 -0.15 -0.01  0.14 -0.05  0.16
0.32  0.35
x9   0.26 -0.11 -0.01 -0.01  0.07  0.06 -0.08  0.07  0.17  1.00 -0.45  0.01 -0.10  0.12  0.02
0.01  0.08
x10 -0.20 -0.23  0.04 -0.31 -0.10 -0.08 -0.07  0.08 -0.15 -0.45  1.00 -0.10  0.06  0.13 -0.17
0.04 -0.20
x11  0.14  0.09  0.27 -0.04  0.00  0.24 -0.01 -0.24 -0.01  0.01 -0.10  1.00  0.13 -0.01  0.20
0.07  0.22
x12  0.24 -0.23  0.06 -0.22 -0.12  0.66  0.08  0.00  0.14 -0.10  0.06  0.13  1.00  0.15  0.35
0.49  0.07
x13 -0.10  0.13  0.15 -0.11 -0.08 -0.15 -0.10  0.06 -0.05  0.12  0.13 -0.01  0.15  1.00  0.17
0.06  0.00
x14  0.07  0.03  0.16  0.05 -0.11  0.36 -0.22 -0.03  0.16  0.02 -0.17  0.20  0.35  0.17  1.00
0.37  0.17
x15  0.26  0.01 -0.02 -0.17 -0.15  0.48 -0.03 -0.10  0.32  0.01  0.04  0.07  0.49  0.06  0.37
1.00  0.39
x16  0.48  0.29  0.14  0.15 -0.30  0.33  0.04 -0.17  0.35  0.08 -0.20  0.22  0.07  0.00  0.17
0.39  1.00

n= 82


P
    y      x1     x2     x3     x4     x5     x6     x7     x8     x9     x10    x11    x12
x13    x14    x15    x16
y          0.1055 0.5039 0.0724 0.4404 0.0000 0.1672 0.9945 0.0000 0.0175 0.0709 0.1989 0.027
6 0.3718 0.5392 0.0167 0.0000
x1  0.1055        0.0456 0.5422 0.0678 0.3000 0.8809 0.0000 0.9586 0.3159 0.0372 0.4286 0.040
3 0.2305 0.8016 0.8985 0.0092
x2  0.5039 0.0456        0.0595 0.5964 0.8384 0.8585 0.0575 0.0012 0.9435 0.7543 0.0137 0.572
1 0.1834 0.1470 0.8404 0.2062
x3  0.0724 0.5422 0.0595        0.0610 0.2354 0.0684 0.1646 0.0301 0.9196 0.0041 0.7236 0.049
2 0.3147 0.6260 0.1197 0.1879
x4  0.4404 0.0678 0.5964 0.0610        0.2662 0.7261 0.3960 0.0414 0.5604 0.3506 0.9793 0.285
1 0.4478 0.3255 0.1765 0.0061
x5  0.0000 0.3000 0.8384 0.2354 0.2662        0.4452 0.2762 0.0007 0.6087 0.4763 0.0272 0.000
0 0.1885 0.0010 0.0000 0.0028
```

```
x6  0.1672 0.8809 0.8585 0.0684 0.7261 0.4452        0.6136 0.2758 0.4518 0.5403 0.9032 0.501
5 0.3818 0.0439 0.8081 0.7394
x7  0.9945 0.0000 0.0575 0.1646 0.3960 0.2762 0.6136        0.1207 0.5100 0.4569 0.0299 0.986
5 0.5909 0.8196 0.3572 0.1215
x8  0.0000 0.9586 0.0012 0.0301 0.0414 0.0007 0.2758 0.1207        0.1326 0.1722 0.9230 0.223
6 0.6376 0.1582 0.0036 0.0011
x9  0.0175 0.3159 0.9435 0.9196 0.5604 0.6087 0.4518 0.5100 0.1326        0.0000 0.9575 0.360
0 0.2967 0.8868 0.9326 0.4916
x10 0.0709 0.0372 0.7543 0.0041 0.3506 0.4763 0.5403 0.4569 0.1722 0.0000        0.3617 0.618
2 0.2330 0.1273 0.7191 0.0726
x11 0.1989 0.4286 0.0137 0.7236 0.9793 0.0272 0.9032 0.0299 0.9230 0.9575 0.3617        0.228
5 0.9630 0.0649 0.5422 0.0515
x12 0.0276 0.0403 0.5721 0.0492 0.2851 0.0000 0.5015 0.9865 0.2236 0.3600 0.6182 0.2285
0.1885 0.0011 0.0000 0.5287
x13 0.3718 0.2305 0.1834 0.3147 0.4478 0.1885 0.3818 0.5909 0.6376 0.2967 0.2330 0.9630 0.188
5        0.1334 0.6082 0.9966
x14 0.5392 0.8016 0.1470 0.6260 0.3255 0.0010 0.0439 0.8196 0.1582 0.8868 0.1273 0.0649 0.001
1 0.1334        0.0007 0.1190
x15 0.0167 0.8985 0.8404 0.1197 0.1765 0.0000 0.8081 0.3572 0.0036 0.9326 0.7191 0.5422 0.000
0 0.6082 0.0007        0.0003
x16 0.0000 0.0092 0.2062 0.1879 0.0061 0.0028 0.7394 0.1215 0.0011 0.4916 0.0726 0.0515 0.528
```

Hide

```
# Scatterplot of all variables
pairs(cbind(y,x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12,x13,x14,x15,x16))
```



Bellow we will fit a regression model to the y values of our data (dependent), using all the xi attributes of our data (independent).

Hide

```
fitall <- lm(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10 + x11 + x12 + x13 + x14 + x
15 + x16)
summary(fitall)
```

```
Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 +
    x10 + x11 + x12 + x13 + x14 + x15 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-4.9744 -0.3346  0.0876  0.5429  2.4008

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.244185   0.159213   1.534 0.129957
x1          -0.357020   0.134447  -2.655 0.009950 **
x2          -0.074085   0.107791  -0.687 0.494337
x3           0.060635   0.105963   0.572 0.569139
x4           0.229306   0.089333   2.567 0.012573 *
x5           0.020863   0.011170   1.868 0.066293 .
x6           0.004180   0.003659   1.143 0.257401
x7          -0.009865   0.070021  -0.141 0.888398
x8           0.573951   0.158859   3.613 0.000590 ***
x9           0.002753   0.003362   0.819 0.415831
x10         -0.145125   0.357077  -0.406 0.685765
x11          0.197892   0.282263   0.701 0.485748
x12          0.018352   0.087056   0.211 0.833695
x13          0.008843   0.013623   0.649 0.518537
x14         -0.597458   0.561641  -1.064 0.291368
x15         -0.014423   0.017679  -0.816 0.417562
x16          0.340328   0.083050   4.098 0.000118 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9945 on 65 degrees of freedom
Multiple R-squared:  0.645, Adjusted R-squared:  0.5576
F-statistic: 7.381 on 16 and 65 DF,  p-value: 2.143e-09
```

Hide

```
AIC(fitall)
```

```
[1] 248.7493
```

Hide

```
BIC(fitall)
```

```
[1] 292.0703
```

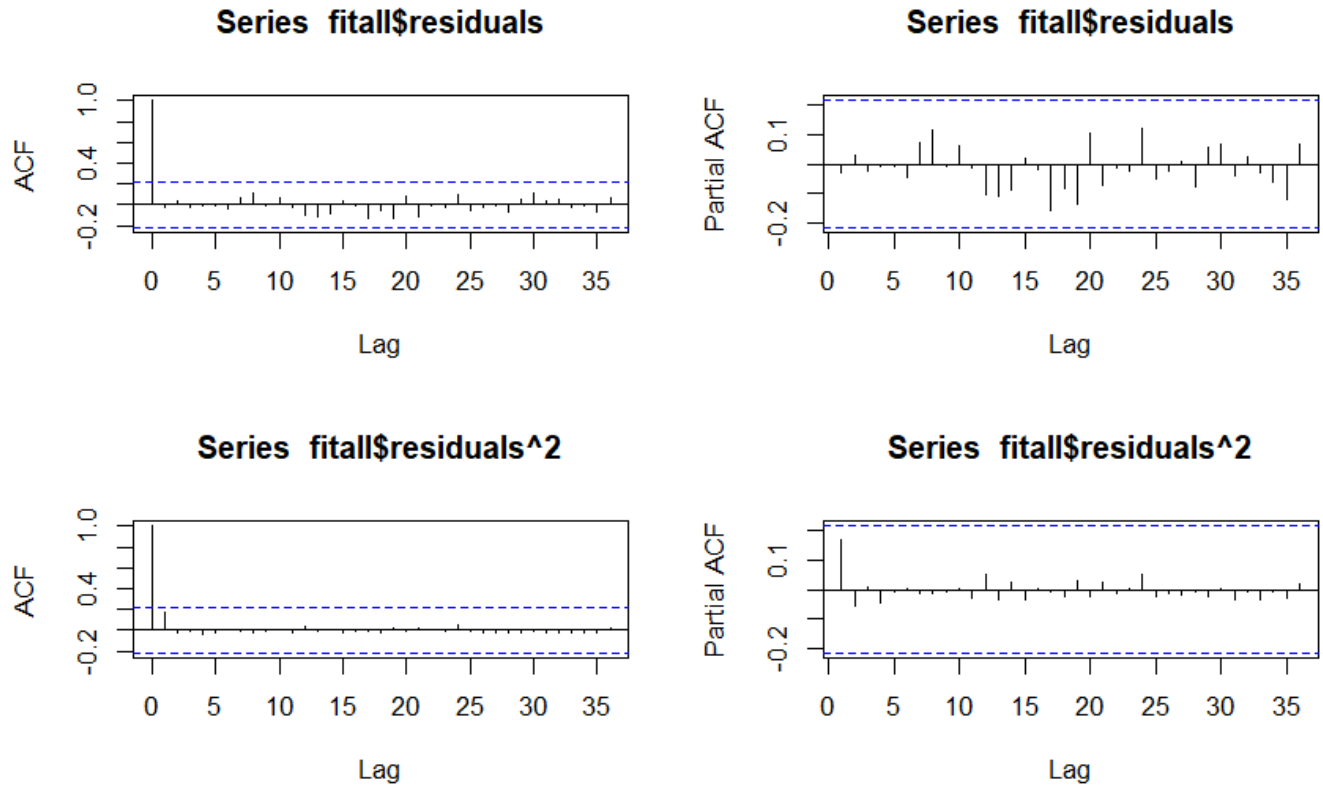
The summary shows us that there are some significant values for estimating y. Our R^2 is 0.645, AIC is 248 and BIC is 292. We will test bellow if the residuals of this model are uncorrelated, homoscedastic and normal.

Hide

```
par(mfrow=c(2,2))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall$residuals, 36)
pacf(fitall$residuals, 36)
```

Hide

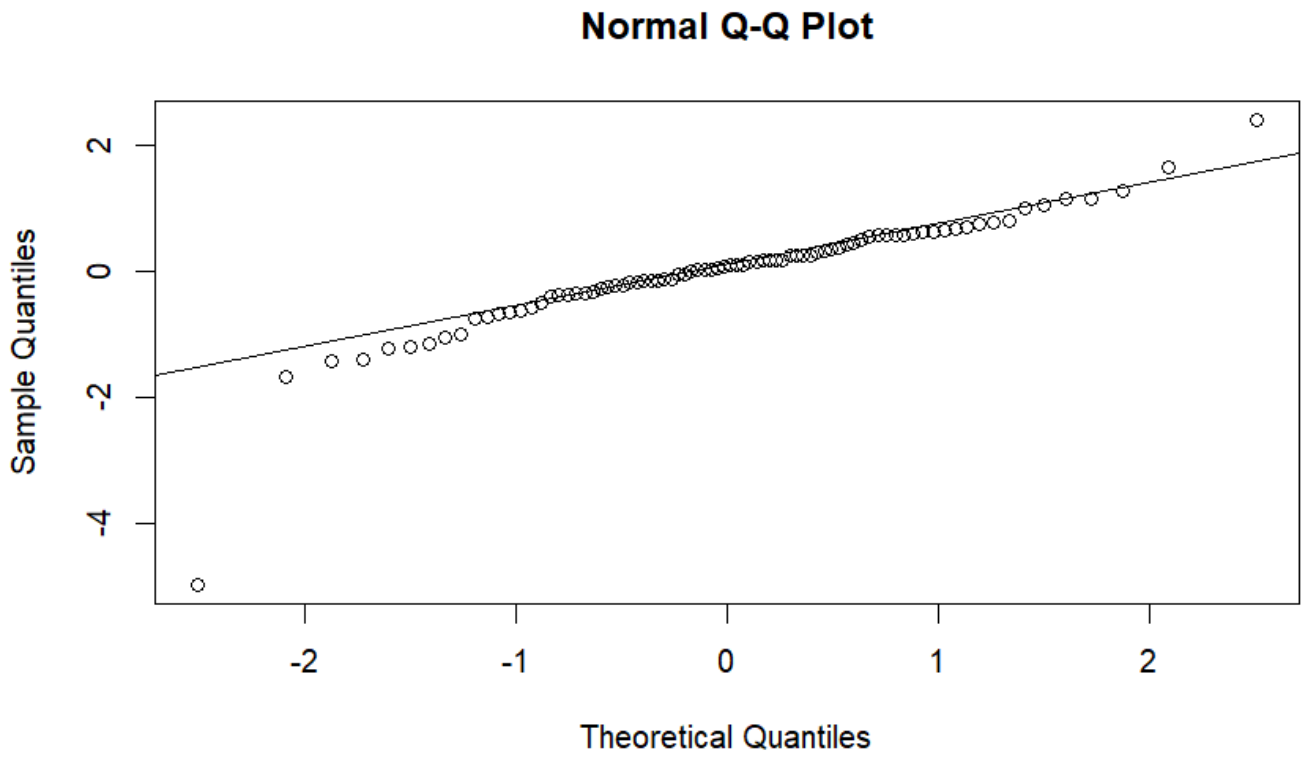
```
# Autocorrelation of the squared residuals
acf(fitall$residuals^2, 36)
pacf(fitall$residuals^2, 36)
```



We can see that our residuals are both uncorrelated and homoscedastic as ACF and PACF aplots stay between the limits, for the residuals and squared residuals.

Hide

```
qqnorm(fitall$residuals)
qqline(fitall$residuals)
```



We can see that for a few tail quantiles the distribution is not normal, but for most it is so we will accept normality for now.

Hide

```
jarque.bera.test(fitall$residuals)
```

```
    Jarque Bera Test

data:  fitall$residuals
X-squared = 443.43, df = 2, p-value < 2.2e-16
```

Hide

```
shapiro.test(fitall$residuals)
```

```
    Shapiro-Wilk normality test

data:  fitall$residuals
W = 0.85022, p-value = 1.252e-07
```

Now we will begin repetitive task to improve our model. We will be fitting the regression model, removing the non statistically significant xi that has the highest p-value each time. We will plot the residuals' and squared residuals' ACF, PACF and normal QQ line plot to check if our assumptions are correct. We will repeat this modeling until AIC or BIC stop improving, all the coefficients are significant of there is a problem with the assumptions on the residuals that can be fixed with time series modeling.

Hide

```
fitall2 <- lm(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x8 + x9 + x10 + x11 + x12 + x13 + x14 + x15 + x16)
summary(fitall2)
```

```

Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x8 + x9 + x10 +
    x11 + x12 + x13 + x14 + x15 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-4.9717 -0.3310  0.0875  0.5460  2.3852

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.239570   0.154646   1.549 0.126126
x1          -0.346943   0.112996  -3.070 0.003102 **
x2          -0.074273   0.106979  -0.694 0.489945
x3           0.059400   0.104812   0.567 0.572820
x4           0.228455   0.088464   2.582 0.012035 *
x5           0.020828   0.011084   1.879 0.064641 .
x6           0.004110   0.003597   1.143 0.257374
x8           0.579008   0.153597   3.770 0.000351 ***
x9           0.002779   0.003332   0.834 0.407322
x10         -0.140575   0.352963  -0.398 0.691715
x11          0.208254   0.270482   0.770 0.444086
x12          0.019962   0.085660   0.233 0.816456
x13          0.008441   0.013221   0.638 0.525396
x14         -0.604069   0.555505  -1.087 0.280805
x15         -0.014444   0.017547  -0.823 0.413353
x16          0.338992   0.081891   4.140 0.000101 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9871 on 66 degrees of freedom
Multiple R-squared:  0.6449,    Adjusted R-squared:  0.5642
F-statistic:  7.99 on 15 and 66 DF,  p-value: 7.413e-10
```

Hide

```
AIC(fitall2)
```

```
[1] 246.7744
```

Hide

```
BIC(fitall2)
```

```
[1] 287.6886
```

Hide

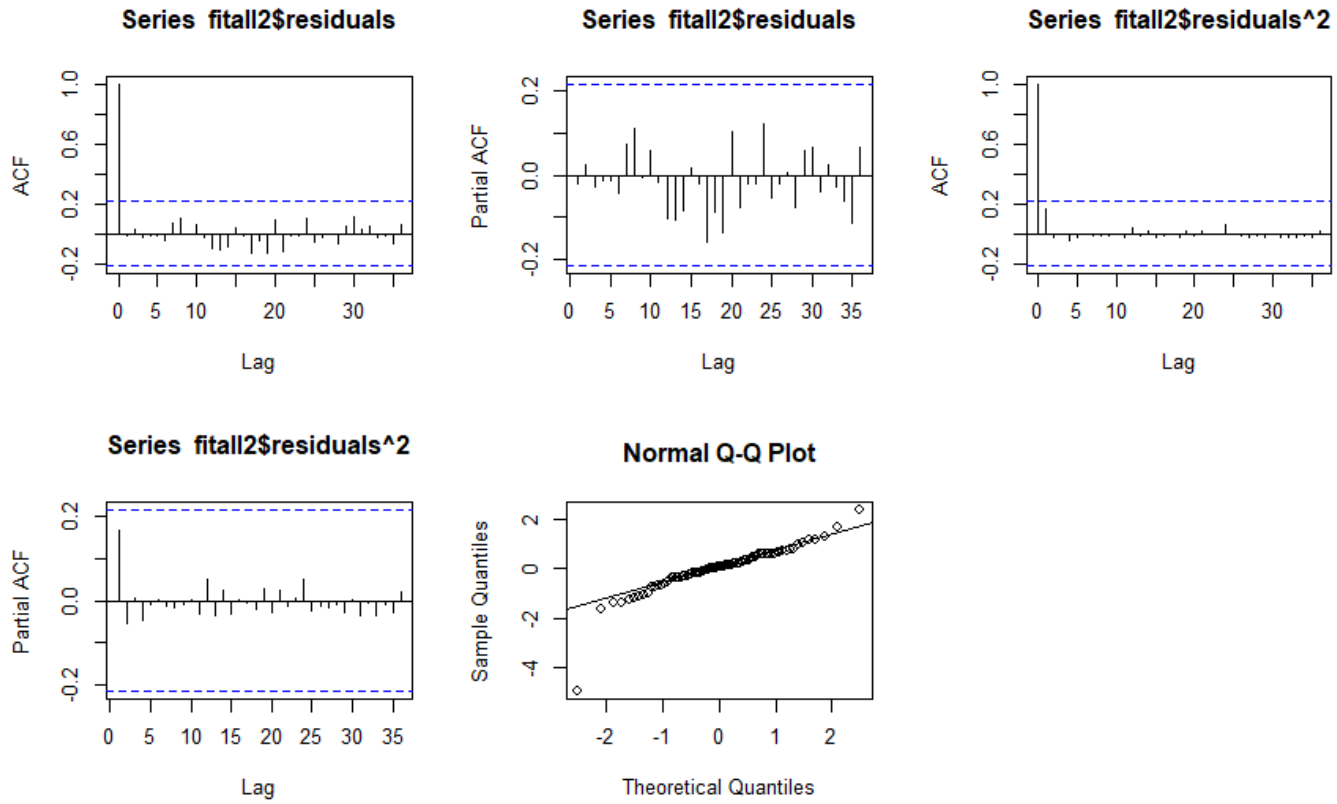
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall2$residuals, 36)
pacf(fitall2$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall2$residuals^2, 36)
pacf(fitall2$residuals^2, 36)
```

Hide

```
qqnorm(fitall2$residuals)
qqline(fitall2$residuals)
```



Hide

```
# Removing x12
fitall3 <- lm(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x8 + x9 + x10 + x11 + x13 + x14 + x15 + x16)
summary(fitall3)
```

```
Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x8 + x9 + x10 +
    x11 + x13 + x14 + x15 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-4.9905 -0.3172  0.0658  0.5271  2.3808

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.247487   0.149800   1.652 0.103189
x1          -0.356526   0.104502  -3.412 0.001099 **
x2          -0.079305   0.104036  -0.762 0.448566
x3           0.053336   0.100812   0.529 0.598506
x4           0.228158   0.087828   2.598 0.011526 *
x5           0.022443   0.008588   2.613 0.011062 *
x6           0.004389   0.003367   1.304 0.196831
x8           0.571934   0.149501   3.826 0.000289 ***
x9           0.002522   0.003123   0.808 0.422123
x10         -0.159676   0.340883  -0.468 0.641006
x11          0.209567   0.268508   0.780 0.437854
x13          0.009807   0.011766   0.834 0.407515
x14         -0.584342   0.545129  -1.072 0.287596
x15         -0.013378   0.016819  -0.795 0.429191
x16          0.337542   0.081076   4.163 9.16e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9801 on 67 degrees of freedom
Multiple R-squared:  0.6446,    Adjusted R-squared:  0.5703
F-statistic: 8.679 on 14 and 67 DF,  p-value: 2.497e-10
```

Hide

```
AIC(fitall3)
```

```
[1] 244.8418
```

Hide

```
BIC(fitall3)
```

```
[1] 283.3493
```

Hide

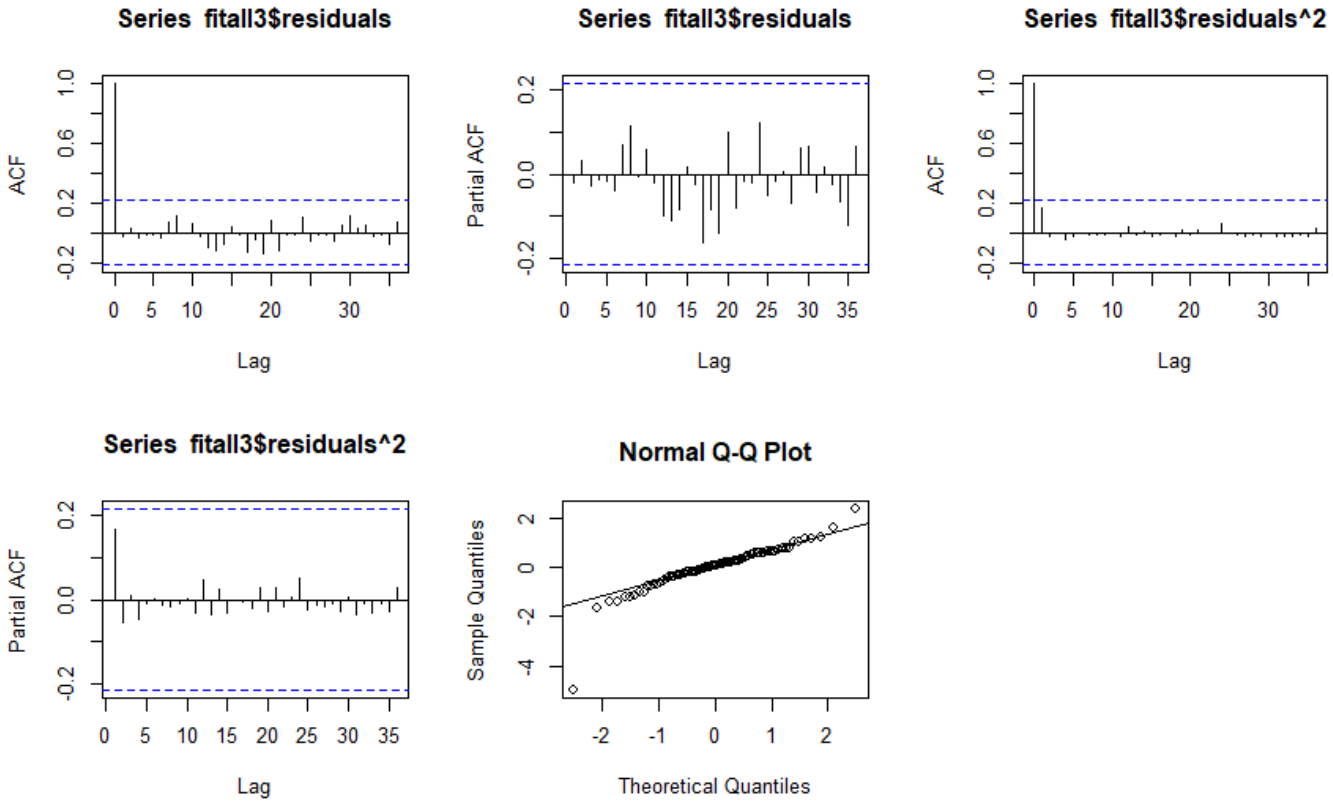
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall3$residuals, 36)
pacf(fitall3$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall3$residuals^2, 36)
pacf(fitall3$residuals^2, 36)
```

Hide

```
qqnorm(fitall3$residuals)
qqline(fitall3$residuals)
```



Hide

```
# Removing x10
fitall4 <- lm(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x8 + x9 + x11 + x13 + x14 + x15 + x16)
summary(fitall4)
```

```
Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x8 + x9 + x11 +
    x13 + x14 + x15 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-4.9747 -0.3210  0.0515  0.5134  2.3903

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.233491   0.145945   1.600 0.114265
x1          -0.337939   0.096119  -3.516 0.000785 ***
x2          -0.072251   0.102348  -0.706 0.482636
x3           0.071250   0.092739   0.768 0.444978
x4           0.231487   0.087036   2.660 0.009746 **
x5           0.022888   0.008486   2.697 0.008811 **
x6           0.004521   0.003336   1.355 0.179866
x8           0.571053   0.148629   3.842 0.000271 ***
x9           0.003362   0.002542   1.322 0.190434
x11          0.208719   0.266957   0.782 0.437019
x13          0.008291   0.011247   0.737 0.463544
x14         -0.538980   0.533370  -1.011 0.315829
x15         -0.013801   0.016698  -0.827 0.411410
x16          0.336009   0.080544   4.172 8.77e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9744 on 68 degrees of freedom
Multiple R-squared:  0.6434,    Adjusted R-squared:  0.5752
F-statistic: 9.438 on 13 and 68 DF,  p-value: 8.686e-11
```

Hide

```
AIC(fitall4)
```

```
[1] 243.1099
```

Hide

```
BIC(fitall4)
```

```
[1] 279.2107
```

Hide

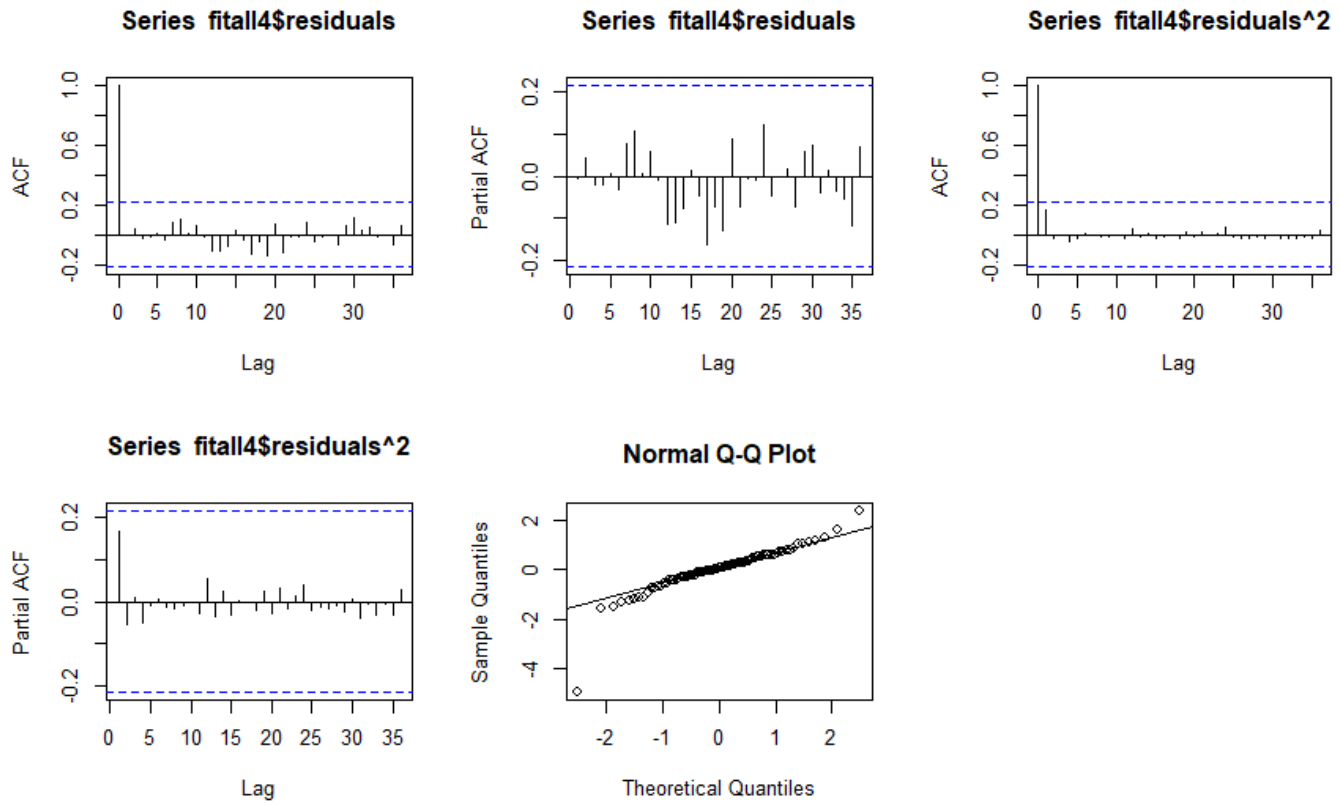
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall4$residuals, 36)
pacf(fitall4$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall4$residuals^2, 36)
pacf(fitall4$residuals^2, 36)
```

Hide

```
qqnorm(fitall4$residuals)
qqline(fitall4$residuals)
```



Hide

```
# Removing x2
fitall5 <- lm(y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x11 + x13 + x14 + x15 + x16)
summary(fitall5)
```

```
Call:
lm(formula = y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x11 + x13 +
    x14 + x15 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-4.9518 -0.3063  0.0451  0.4925  2.6210

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.182070   0.126007   1.445 0.153007
x1          -0.307965   0.085918  -3.584 0.000626 ***
x3           0.090125   0.088478   1.019 0.311947
x4           0.229277   0.086663   2.646 0.010091 *
x5           0.023306   0.008435   2.763 0.007332 **
x6           0.004147   0.003282   1.264 0.210642
x8           0.613628   0.135348   4.534 2.37e-05 ***
x9           0.003440   0.002531   1.360 0.178414
x11          0.167393   0.259510   0.645 0.521045
x13          0.006958   0.011047   0.630 0.530858
x14         -0.635309   0.513742  -1.237 0.220417
x15         -0.011627   0.016352  -0.711 0.479451
x16          0.311745   0.072576   4.295 5.59e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9709 on 69 degrees of freedom
Multiple R-squared:  0.6408,    Adjusted R-squared:  0.5783
F-statistic: 10.26 on 12 and 69 DF,  p-value: 3.289e-11
```

Hide

```
AIC(fitall5)
```

```
[1] 241.7087
```

Hide

```
BIC(fitall5)
```

```
[1] 275.4027
```

Hide

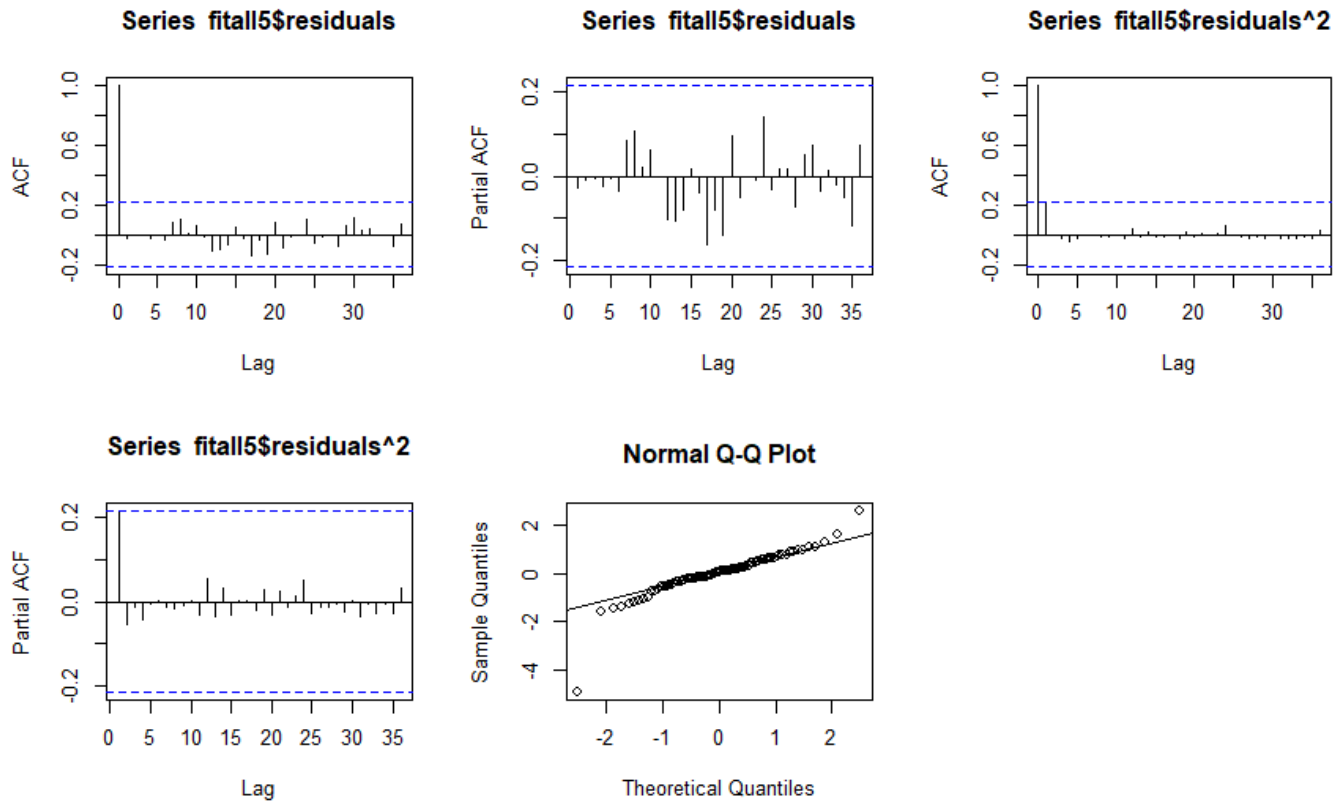
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall5$residuals, 36)
pacf(fitall5$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall5$residuals^2, 36)
pacf(fitall5$residuals^2, 36)
```

Hide

```
qqnorm(fitall5$residuals)
qqline(fitall5$residuals)
```



Hide

```
# Removing x13
fitall6 <- lm(y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x11 + x14 + x15 + x16)
summary(fitall6)
```

```
Call:
lm(formula = y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x11 + x14 +
    x15 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-5.0251 -0.2843  0.0007  0.4947  2.5967

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.190682   0.124722   1.529 0.130808
x1          -0.300104   0.084640  -3.546 0.000703 ***
x3           0.082285   0.087220   0.943 0.348710
x4           0.219970   0.085025   2.587 0.011758 *
x5           0.022125   0.008188   2.702 0.008637 **
x6           0.004083   0.003266   1.250 0.215342
x8           0.612201   0.134745   4.543 2.25e-05 ***
x9           0.003704   0.002485   1.490 0.140590
x11          0.167912   0.258388   0.650 0.517920
x14         -0.569703   0.500901  -1.137 0.259267
x15         -0.010936   0.016244  -0.673 0.503028
x16          0.309385   0.072166   4.287 5.68e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9667 on 70 degrees of freedom
Multiple R-squared:  0.6387,    Adjusted R-squared:  0.582
F-statistic: 11.25 on 11 and 70 DF,  p-value: 1.135e-11
```

Hide

```
AIC(fitall6)
```

```
[1] 240.1788
```

Hide

```
BIC(fitall6)
```

```
[1] 271.4662
```

Hide

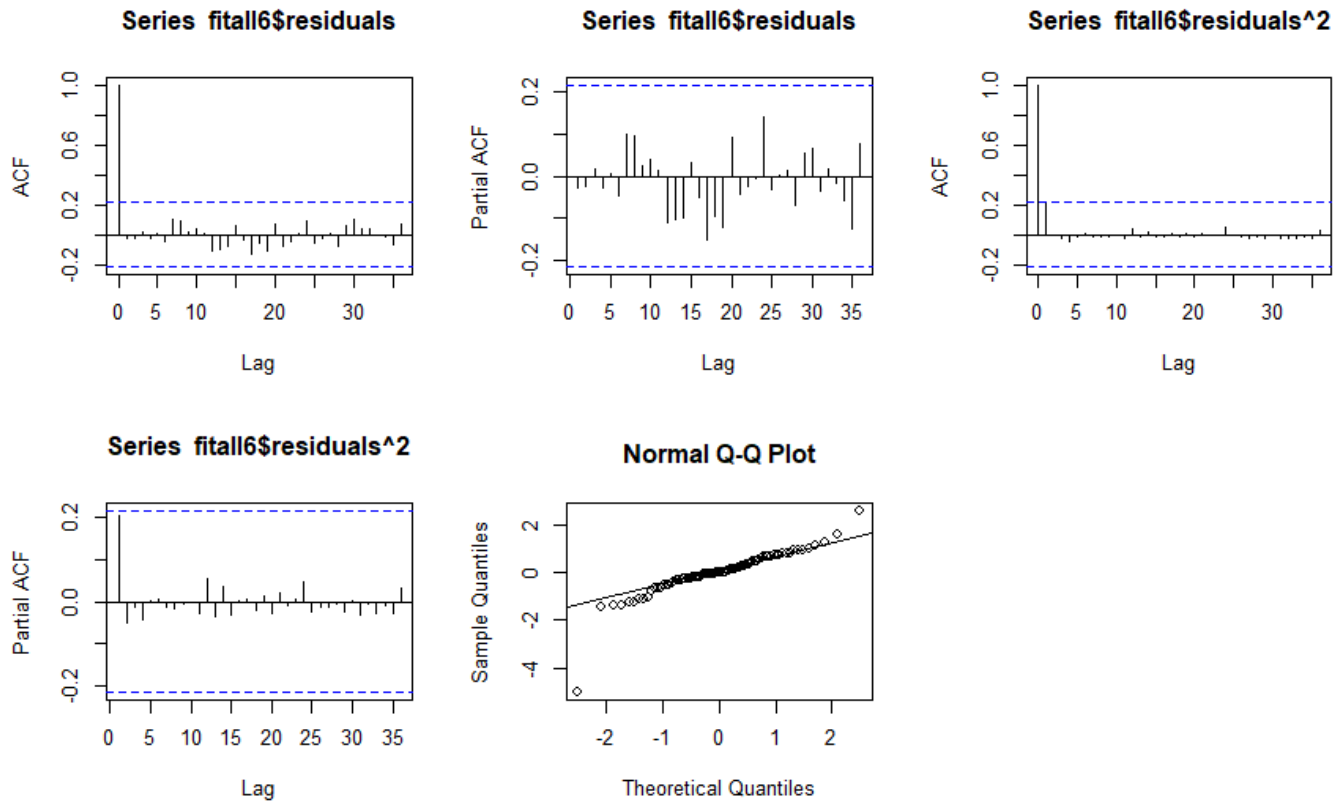
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall6$residuals, 36)
pacf(fitall6$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall6$residuals^2, 36)
pacf(fitall6$residuals^2, 36)
```

Hide

```
qqnorm(fitall6$residuals)
qqline(fitall6$residuals)
```



Hide

```
# Removing x15
fitall7 <- lm(y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x11 + x14 + x16)
summary(fitall7)
```

```
Call:
lm(formula = y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x11 + x14 +
    x16)

Residuals:
    Min      1Q  Median      3Q     Max
-5.0924 -0.2864 -0.0052  0.4600  2.5941

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.178378   0.122900   1.451 0.151070
x1          -0.295814   0.084074  -3.519 0.000762 ***
x3           0.099324   0.083145   1.195 0.236223
x4           0.220205   0.084697   2.600 0.011335 *
x5           0.020870   0.007942   2.628 0.010525 *
x6           0.003898   0.003242   1.202 0.233214
x8           0.597883   0.132542   4.511 2.49e-05 ***
x9           0.003827   0.002469   1.550 0.125520
x11          0.190480   0.255216   0.746 0.457920
x14         -0.668367   0.477133  -1.401 0.165630
x16          0.295321   0.068810   4.292 5.51e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.963 on 71 degrees of freedom
Multiple R-squared:  0.6364,    Adjusted R-squared:  0.5852
F-statistic: 12.43 on 10 and 71 DF,  p-value: 3.827e-12
```

Hide

```
AIC(fitall7)
```

```
[1] 238.708
```

Hide

```
BIC(fitall7)
```

```
[1] 267.5886
```

Hide

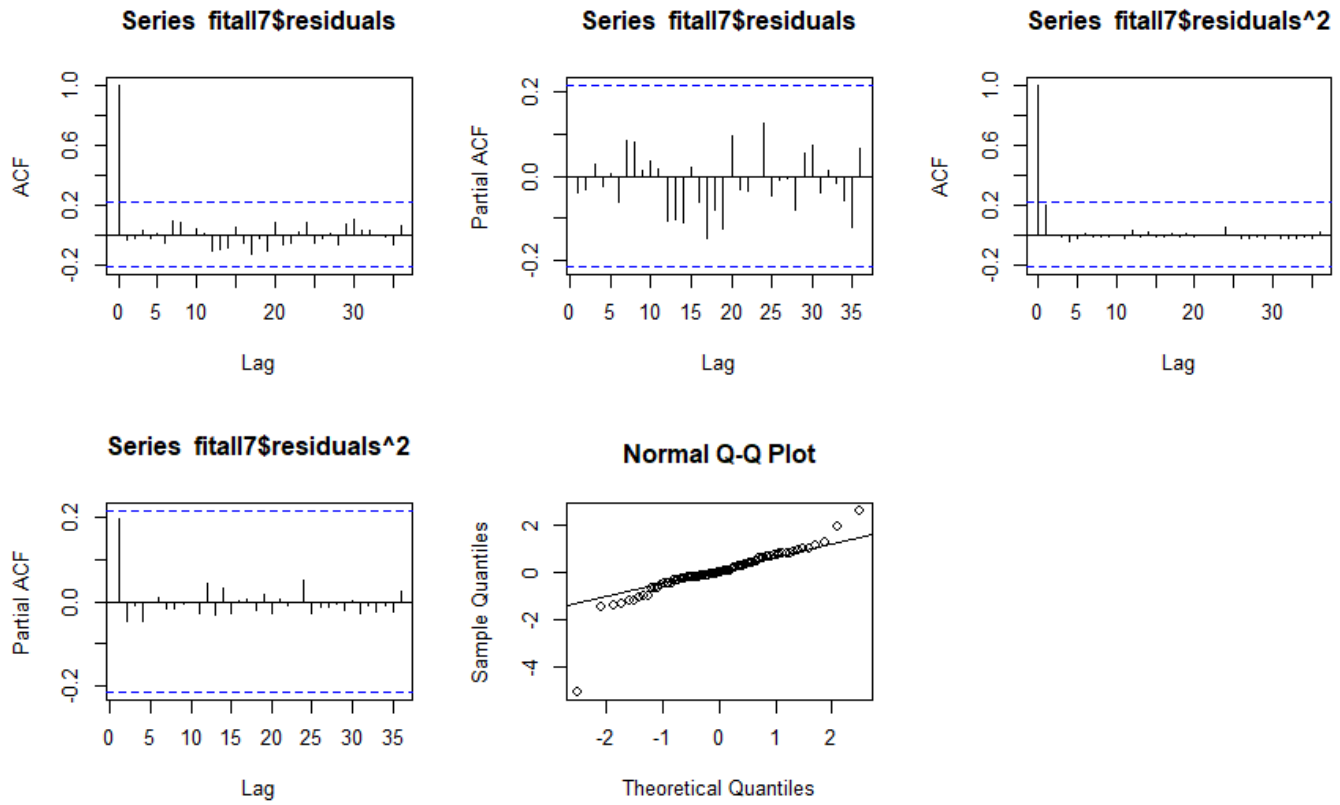
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall7$residuals, 36)
pacf(fitall7$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall7$residuals^2, 36)
pacf(fitall7$residuals^2, 36)
```

Hide

```
qqnorm(fitall7$residuals)
qqline(fitall7$residuals)
```



Hide

```
# Removing x11
fitall8 <- lm(y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x14 + x16)
summary(fitall8)
```

```
Call:
lm(formula = y ~ x1 + x3 + x4 + x5 + x6 + x8 + x9 + x14 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-5.0759 -0.3093 -0.0132  0.4765  2.5941

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.177354   0.122513   1.448 0.152061
x1          -0.293016   0.083731  -3.499 0.000804 ***
x3           0.099289   0.082889   1.198 0.234901
x4           0.223272   0.084336   2.647 0.009958 **
x5           0.021933   0.007789   2.816 0.006273 **
x6           0.004025   0.003227   1.247 0.216321
x8           0.582721   0.130573   4.463 2.93e-05 ***
x9           0.003852   0.002461   1.565 0.121943
x14         -0.620916   0.471421  -1.317 0.191976
x16          0.303134   0.067800   4.471 2.84e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.96 on 72 degrees of freedom
Multiple R-squared:  0.6335,    Adjusted R-squared:  0.5877
F-statistic: 13.83 on 9 and 72 DF,  p-value: 1.282e-12
```

Hide

```
AIC(fitall8)
```

```
[1] 237.3488
```

Hide

```
BIC(fitall8)
```

```
[1] 263.8227
```

Hide

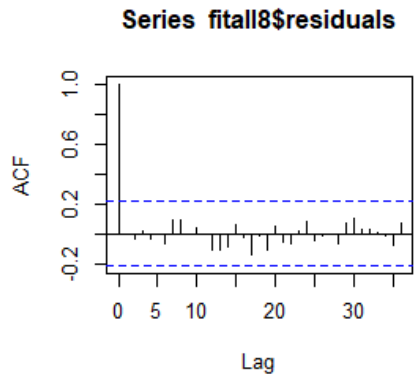
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall8$residuals, 36)
pacf(fitall8$residuals, 36)
```

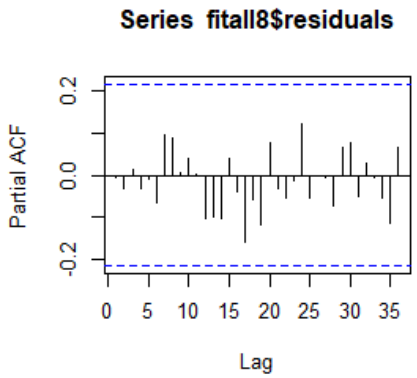
Hide

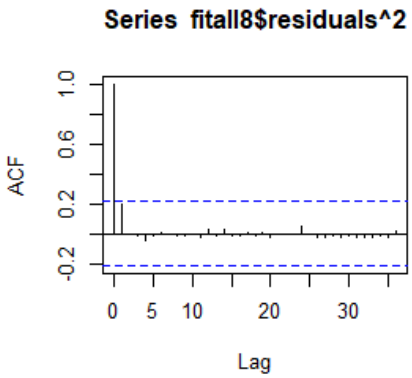
```
# Autocorrelation of the squared residuals
acf(fitall8$residuals^2, 36)
pacf(fitall8$residuals^2, 36)
```

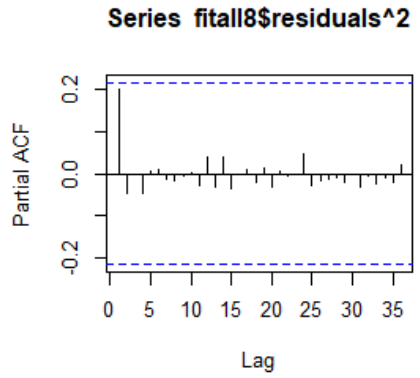
Hide

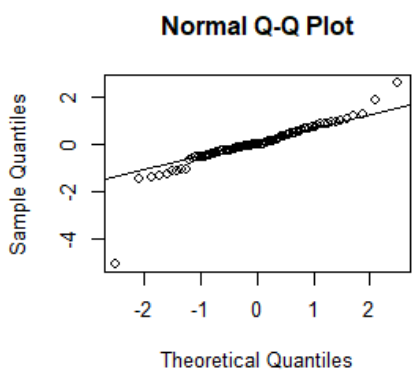
```
qqnorm(fitall8$residuals)
qqline(fitall8$residuals)
```











Hide

```
# Removing x3
fitall9 <- lm(y ~ x1 + x4 + x5 + x6 + x8 + x9 + x14 + x16)
summary(fitall9)
```

```
Call:
lm(formula = y ~ x1 + x4 + x5 + x6 + x8 + x9 + x14 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-5.0583 -0.3767 -0.0034  0.5410  2.5634

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.216380   0.118453   1.827 0.071832 .
x1          -0.307086   0.083150  -3.693 0.000425 ***
x4           0.214202   0.084246   2.543 0.013123 *
x5           0.019129   0.007451   2.567 0.012301 *
x6           0.004684   0.003189   1.469 0.146218
x8           0.620152   0.127155   4.877 6.11e-06 ***
x9           0.003708   0.002465   1.504 0.136859
x14         -0.542208   0.468208  -1.158 0.250619
x16          0.314222   0.067364   4.665 1.37e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9629 on 73 degrees of freedom
Multiple R-squared:  0.6262,    Adjusted R-squared:  0.5853
F-statistic: 15.29 on 8 and 73 DF,  p-value: 6.145e-13
```

Hide

```
AIC(fitall9)
```

```
[1] 236.9669
```

Hide

```
BIC(fitall9)
```

```
[1] 261.0341
```

Hide

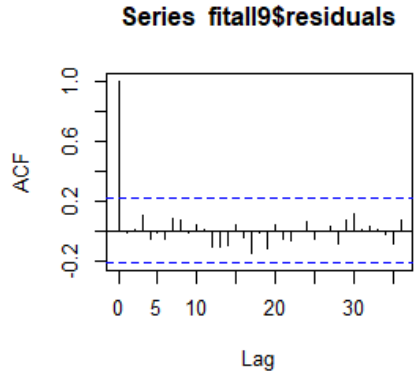
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall9$residuals, 36)
pacf(fitall9$residuals, 36)
```

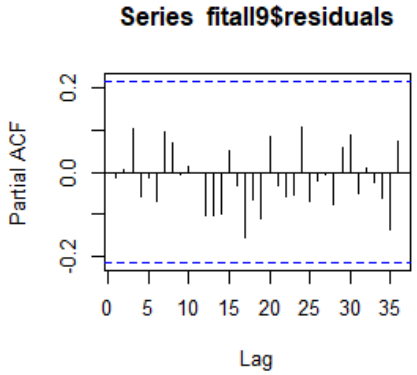
Hide

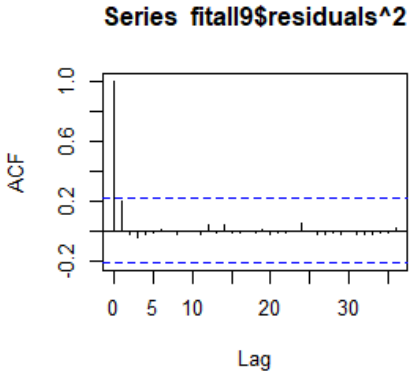
```
# Autocorrelation of the squared residuals
acf(fitall9$residuals^2, 36)
pacf(fitall9$residuals^2, 36)
```

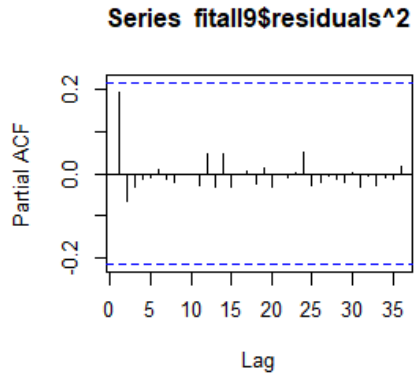
Hide

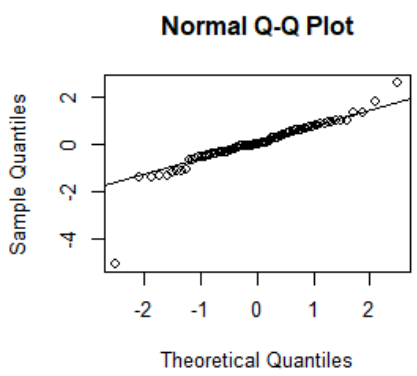
```
qqnorm(fitall9$residuals)
qqline(fitall9$residuals)
```











Hide

```
# Removing x14
fitall10 <- lm(y ~ x1 + x4 + x5 + x6 + x8 + x9 + x16)
summary(fitall10)
```

```
Call:
lm(formula = y ~ x1 + x4 + x5 + x6 + x8 + x9 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-5.1995 -0.3070 -0.0303  0.5386  2.6420

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.255083   0.113902   2.239 0.028129 *
x1          -0.314188   0.083114  -3.780 0.000315 ***
x4           0.221613   0.084196   2.632 0.010322 *
x5           0.016615   0.007144   2.326 0.022786 *
x6           0.005491   0.003120   1.760 0.082482 .
x8           0.613923   0.127334   4.821 7.42e-06 ***
x9           0.003753   0.002471   1.519 0.133040
x16          0.312974   0.067511   4.636 1.50e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9651 on 74 degrees of freedom
Multiple R-squared:  0.6194,    Adjusted R-squared:  0.5834
F-statistic:  17.2 on 7 and 74 DF,  p-value: 2.659e-13
```

Hide

```
AIC(fitall10)
```

```
[1] 236.4597
```

Hide

```
BIC(fitall10)
```

```
[1] 258.1201
```

Hide

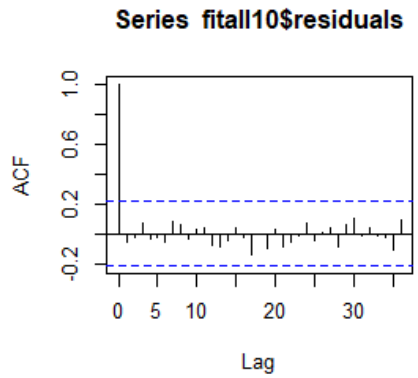
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall10$residuals, 36)
pacf(fitall10$residuals, 36)
```

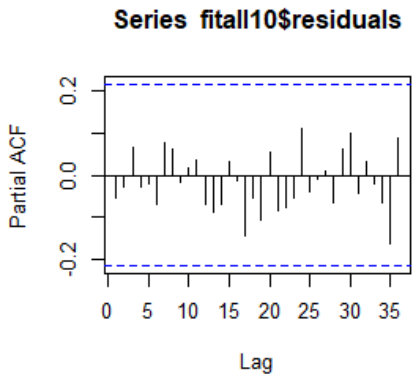
Hide

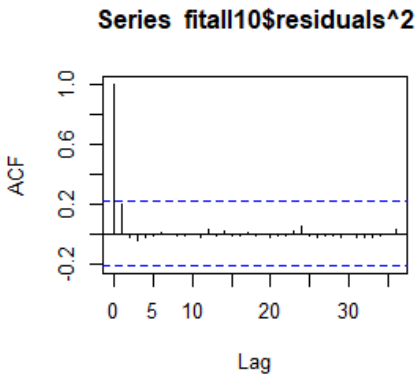
```
# Autocorrelation of the squared residuals
acf(fitall10$residuals^2, 36)
pacf(fitall10$residuals^2, 36)
```

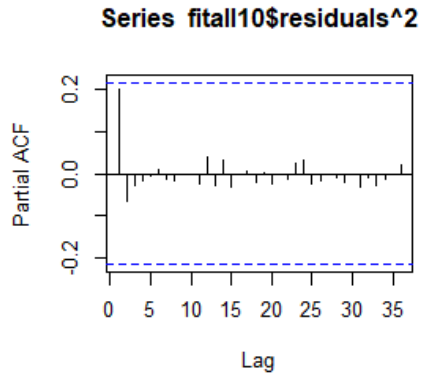
Hide

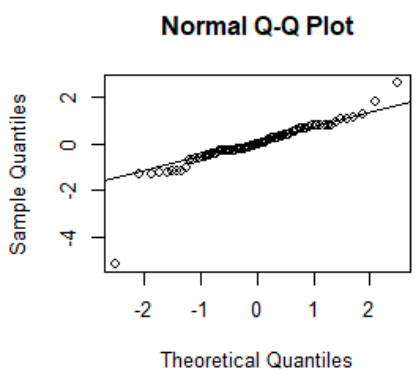
```
qqnorm(fitall10$residuals)
qqline(fitall10$residuals)
```











Hide

```
# Removing x9
fitall11 <- lm(y ~ x1 + x4 + x5 + x6 + x8 + x16)
summary(fitall11)
```

```
Call:
lm(formula = y ~ x1 + x4 + x5 + x6 + x8 + x16)

Residuals:
    Min      1Q  Median      3Q     Max
-5.4016 -0.3848 -0.0262  0.5235  2.6632

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.249966   0.114840   2.177 0.032656 *
x1          -0.336897   0.082468  -4.085 0.000109 ***
x4           0.242931   0.083738   2.901 0.004878 **
x5           0.015896   0.007191   2.211 0.030108 *
x6           0.004962   0.003127   1.587 0.116737
x8           0.648485   0.126371   5.132 2.19e-06 ***
x16          0.325441   0.067591   4.815 7.47e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9735 on 75 degrees of freedom
Multiple R-squared:  0.6075,    Adjusted R-squared:  0.5761
F-statistic: 19.35 on 6 and 75 DF,  p-value: 1.703e-13
```

Hide

```
AIC(fitall11)
```

```
[1] 236.9772
```

Hide

```
BIC(fitall11)
```

```
[1] 256.231
```

Hide

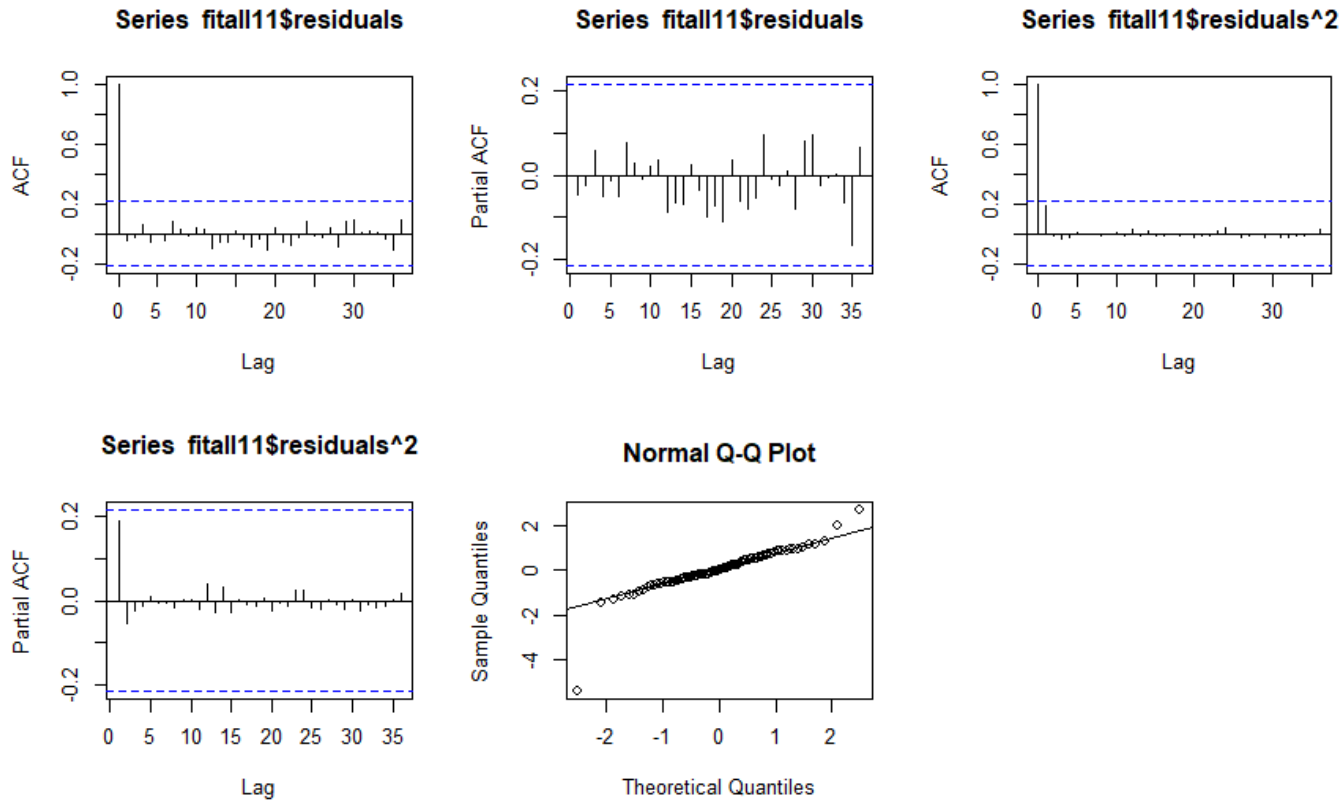
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall11$residuals, 36)
pacf(fitall11$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall11$residuals^2, 36)
pacf(fitall11$residuals^2, 36)
```

Hide

```
qqnorm(fitall11$residuals)
qqline(fitall11$residuals)
```



Hide

```
# Removing x6
fitall12 <- lm(y ~ x1 + x4 + x5 + x8 + x16)
summary(fitall12)
```

```

Call:
lm(formula = y ~ x1 + x4 + x5 + x8 + x16)

Residuals:
    Min      1Q  Median      3Q     Max 
-5.7233 -0.4566  0.0903  0.4810  2.6430 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.247131   0.115968   2.131 0.036321 *  
x1          -0.337922   0.083285  -4.057 0.000119 ***
x4           0.241633   0.084566   2.857 0.005509 ** 
x5           0.014283   0.007189   1.987 0.050554 .  
x8           0.678528   0.126187   5.377 8.05e-07 ***
x16          0.327998   0.068244   4.806 7.58e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9831 on 76 degrees of freedom
Multiple R-squared:  0.5943,    Adjusted R-squared:  0.5676 
F-statistic: 22.27 on 5 and 76 DF,  p-value: 1.122e-13
```

Hide

```
AIC(fitall12)
```

```
[1] 237.6854
```

Hide

```
BIC(fitall12)
```

```
[1] 254.5324
```

Hide

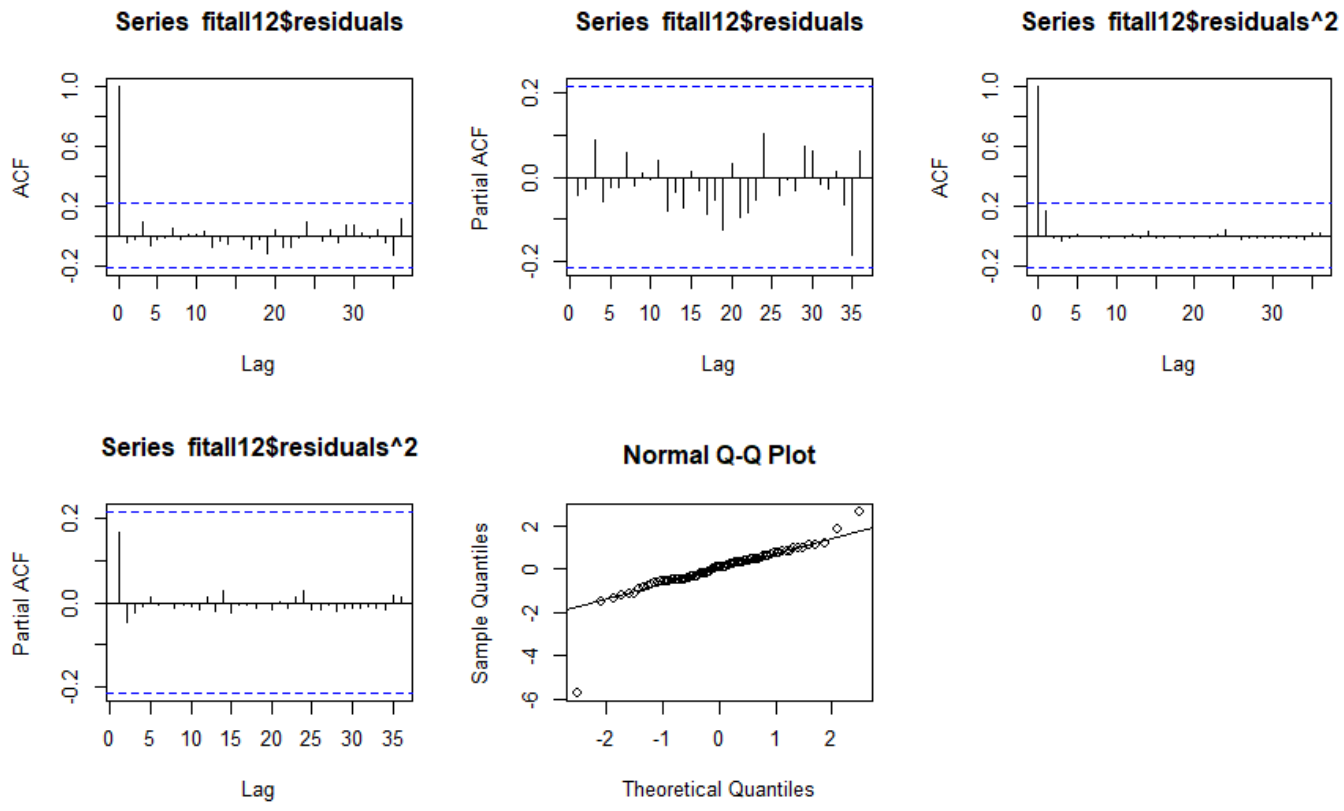
```
par(mfrow=c(2,3))
# Diagnostic tests for the residuals
# Autocorrelation of the residuals
acf(fitall12$residuals, 36)
pacf(fitall12$residuals, 36)
```

Hide

```
# Autocorrelation of the squared residuals
acf(fitall12$residuals^2, 36)
pacf(fitall12$residuals^2, 36)
```

Hide

```
qqnorm(fitall12$residuals)
qqline(fitall12$residuals)
```



Now all the coefficients are significant, The residuals are uncorrelated (ACF, PACF plots of residuals), Homoscedastic (ACF, PACF plots of squared residuals) and close to normal (a few quantiles not on normal lines)

Hide

```
Box.test(fitall12$residuals,lag=12,type="Ljung")
```

```
    Box-Ljung test

data:  fitall12$residuals
X-squared = 2.3516, df = 12, p-value = 0.9986
```

Hide

```
Box.test(fitall12$residuals^2,lag=12,type="Ljung")
```

```
    Box-Ljung test

data:  fitall12$residuals^2
X-squared = 2.6547, df = 12, p-value = 0.9975
```

Hide

```
jarque.bera.test(fitall12$residuals)
```

```
    Jarque Bera Test

data:  fitall12$residuals
X-squared = 810.4, df = 2, p-value < 2.2e-16
```

Hide

```
shapiro.test(fitall12$residuals)
```

```
    Shapiro-Wilk normality test

data:  fitall12$residuals
W = 0.80786, p-value = 5.856e-09
```

Our final model is y = 0.247131 - 0.337922*x1 + 0.241633*x4 + 0.014283 *x5 + 0.678528*x8 + 0.327998*x16. It translates to RGDPGrowth = 0.247131 -0.337922*(LAG1 NED) + 0.241633*(LAG4 NED) + 0.014283* (OIL WTI) + 0.678528*(RPROD NED) + 0.327998*(LEADNED) with R-squared: 0.5676, AIC: 237, BIC: 254.